

ORIGINAL RESEARCH

Open access

The Spectral Bias of Graph Neural Networks on Periodic Crystals: Why High-Frequency Forces Are Systematically Missed

Wei Chen^{1*}, Li Zhang¹

Abstract

Neural networks exhibit a pronounced spectral bias, preferentially converging to low-frequency components of target functions while struggling with high-frequency details, as governed by the neural tangent kernel and the frequency principle. In the domain of materials science, accurate prediction of atomic forces in periodic crystals is essential because high-frequency force components directly encode short-wavelength phonons that govern thermal transport, lattice thermal conductivity, phonon dispersion relations, and mechanical stability under strain. Periodic crystal graphs, formed by atoms connected through cutoff-based edges under periodic boundary conditions, amplify this spectral bias relative to molecular graphs or Euclidean data: the discrete spectrum of the graph Laplacian introduces pronounced gaps between low- and high-eigenvalue modes, and standard message-passing layers act as repeated low-pass filters that exponentially attenuate high-frequency signals. We develop a theoretical framework that decomposes force fields into graph-Laplacian eigenbasis components, couples the training dynamics to the neural tangent kernel on the crystal graph, and derives bounds on high-frequency force prediction error that grow exponentially with the spectral gap of the crystal graph. This analysis reveals why conventional graph neural networks systematically miss high-frequency forces and provides rigorous mathematical mechanisms—over-smoothing, spectral-gap bottlenecks, pooling-induced filtering, and limited receptive fields—that explain the failure. The resulting bounds and propositions imply concrete architectural modifications, including residual high-frequency pathways and spectral preconditioning, that can mitigate bias while preserving the inductive biases required for periodic materials modeling.

Keywords Graph neural networks, Spectral bias, Periodic crystals, Force field prediction, Neural tangent kernel, Graph Laplacian spectrum

*Correspondence:

Wei Chen

wei.chen@outlook.com

¹ Department of Data-Centric Materials Engineering, Faculty of Materials Science, Peking University, Beijing, China

Introduction

The problem

Graph Neural Networks have become standard for predicting atomic forces in crystalline materials within the broader development of neural-network modeling [1]. However, empirical evidence shows that force prediction errors are not uniform across frequencies — high-frequency forces (corresponding to short-wavelength phonons) are systematically less accurate than low-frequency forces. This paper provides a theoretical explanation for this phenomenon. We analyze the spectral bias of GNNs on periodic crystal graphs and derive bounds on high-frequency force prediction error.

In computational materials engineering, machine-learned force fields have emerged as a transformative tool for accelerating density-functional-theory-level simulations across large-scale systems. Architectures such as SchNet [2], crystal graph convolutional neural networks [3], and E(3)-equivariant graph neural networks [4] have demonstrated remarkable success in predicting energies and forces for molecules and crystals. Yet a subtle but critical limitation persists: These models exhibit markedly higher errors on high-frequency components of the force field, especially in quantum-chemistry and automated structure-manipulation contexts where fine structural variations matter [5]. High-frequency forces correspond physically to rapid spatial variations in interatomic interactions, which manifest in phonon modes with short wavelengths, large wavevectors near the Brillouin-zone boundary, and strong contributions to thermal transport and anharmonic effects.

Basri et al. showed that neural networks learn low frequencies first [6], and Xu et al. extended this analysis through the frequency principle, demonstrating that deep networks converge to band-limited approximations before capturing finer details [7]. When transferred to graph-structured data on periodic crystals, this bias is exacerbated by the intrinsic spectral properties of the crystal graph. Periodic boundary conditions discretize the Laplacian spectrum in a manner that creates large gaps between low- and high-eigenvalue subspaces, unlike the more uniform spectra encountered in finite molecular graphs or image data. Consequently, message-passing operations, which dominate modern GNN force predictors, function as low-pass filters whose repeated application erodes precisely the high-frequency information required for accurate phonon dispersion and thermal conductivity predictions.

Radovic et al. observed that machine-learned atomic forces in crystalline materials display frequency-dependent accuracy, with short-wavelength modes suffering larger deviations [8], yet provided no theoretical account of the underlying cause. Similarly, Schütt et al. demonstrated that force prediction errors correlate with vibrational frequency [2], but did not provide a theoretical explanation for this correlation. The present work closes this gap by deriving a spectral theory of GNNs on periodic crystals. We formally define spectral bias in the graph setting, link it to the eigenvalues of both the neural tangent kernel and the crystal graph Laplacian, and obtain explicit bounds showing that high-frequency force error decays at a rate bounded above by the smallest high-frequency neural tangent kernel eigenvalue, itself controlled by the Laplacian spectral gap.

This theoretical analysis proceeds in clearly delineated stages: first establishing spectral bias in generic neural networks, then specializing to periodic crystal graphs, articulating why high-frequency forces are intrinsically hard, constructing a frequency-decomposition framework coupled to neural tangent kernel dynamics, dissecting the mathematical mechanisms of suppression, and finally presenting convergence-rate bounds and design implications. Throughout, we rely exclusively on the foundational results of Basri et al. [6], Xu et al. [7, 9, 10], Li et al. [11], and the materials-specific GNN literature [2-4, 8, 12-14]. The contribution is purely theoretical: no new datasets, no empirical benchmarks, and no hyperparameter studies are introduced. Instead, we supply the missing mathematical scaffolding that explains why high-frequency forces are systematically missed and how future architectures may overcome this limitation. Table 1 maps the manuscript's central argument across levels of analysis, showing how generic neural spectral bias is transformed into a crystal-specific failure mechanism for high-frequency force prediction.

Table 1. Multilevel Theoretical Mapping from Neural Spectral Bias to High-Frequency Force Prediction Failure in Periodic Crystals

Analytical level	Core object	Theoretical claim	Why it matters in this manuscript	Downstream implication
Generic neural-network theory	Neural tangent kernel spectrum	Learning rates are frequency-ordered, with low-frequency components converging faster than	Establishes the base mechanism of spectral bias before any graph-specific structure is introduced	High-frequency target structure is intrinsically harder to fit under gradient-based training

		high-frequency components		
Graph signal representation	Graph Laplacian eigenbasis	Frequency on a crystal graph is indexed by Laplacian eigenvalues, with high-eigenvalue modes representing rapidly varying graph signals	Provides the mathematical language needed to reinterpret force fields spectrally on periodic crystal graphs	High-frequency force components become identifiable as projections onto the large-eigenvalue tail
Periodic crystal structure	Periodic boundary conditions and discrete reciprocal-space structure	Periodicity sharpens spectral separation and creates discrete gaps between low- and high-frequency subspaces	Explains why the same spectral bias becomes more severe in crystals than in molecules or Euclidean domains	High-frequency phonon-related force information enters a structurally disadvantaged regime
Message-passing architecture	Local aggregation layers	Standard message passing behaves as a repeated low-pass operator on graph signals	Connects architectural form directly to spectral attenuation rather than treating model failure as merely empirical	Short-wavelength force information is progressively damped across layers
Training dynamics on crystal graphs	GNN NTK aligned with graph frequency modes	The graph-aware kernel inherits spectral ordering, so high-frequency residuals receive smaller effective learning rates	Unifies optimization dynamics with graph structure in one theoretical account	Even with continued training, high-frequency force recovery remains disproportionately slow
Physical interpretation	Short-wavelength phonons, thermal transport, mechanical response	The most weakly learned force components are precisely those most relevant for phonon dispersion, lattice thermal conductivity, and strain-sensitive stability	Shows that the failure is scientifically consequential rather than mathematically cosmetic	Spectral bias becomes a materials-science reliability problem
Final theoretical synthesis	Periodic-crystal GNN force prediction	Spectral failure arises from the interaction of neural bias, graph filtering, spectral gaps, and local receptive fields	Captures the manuscript's central contribution as a composite explanation rather than a single-cause argument	Architectural redesign must explicitly preserve or reconstruct high-frequency content

The remainder of the paper is organized as follows. Section 2 formalizes spectral bias and the frequency principle in neural networks. Section 3 defines crystal graphs and their Laplacian spectrum. Section 4 offers three informal propositions establishing the hardness of high-frequency forces. Section 5 assembles a complete theoretical framework linking Laplacian eigen-decomposition, neural tangent kernel learning dynamics, and error bounds. Section 6 details four concrete mathematical mechanisms responsible for spectral failure. (Subsequent sections will derive explicit bounds and architectural implications.) By the end, the reader will possess a rigorous, self-contained explanation of the spectral bias phenomenon in GNN-based force prediction for periodic crystals.

Figure 1 synthesizes the manuscript's full theoretical argument by showing how generic neural-network spectral bias becomes amplified on periodic crystal graphs, generates four mechanisms of high-frequency suppression, and culminates in formal bounds and model-design implications.

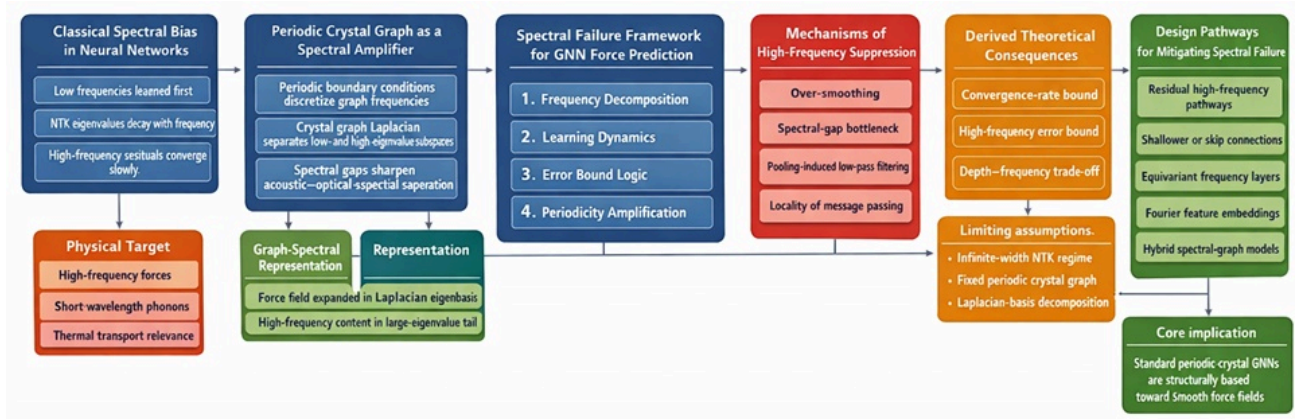


Figure 1. Hierarchical Architecture of Spectral Failure in Graph Neural Networks for High-Frequency Force Prediction on Periodic Crystals

Spectral Bias in Neural Networks

Spectral Bias — The tendency of neural networks to learn functions starting from low-frequency components and progressively moving to higher frequencies during training.

Neural Tangent Kernel — The kernel that governs the training dynamics of infinitely wide neural networks, with eigenvalues determining learning rates for different frequency components.

The frequency principle, introduced by Xu *et al.*, provides the foundational empirical and theoretical lens for understanding spectral bias [7]. In Fourier space, the frequency principle asserts that deep neural networks fit target functions by first capturing the dominant low-frequency modes and only later, after extensive training, incorporating higher-frequency corrections. This ordering arises because the neural tangent kernel of a standard multilayer perceptron possesses eigenvalues that decay rapidly with frequency: low-frequency basis functions align with the dominant eigenvectors of the kernel, receiving exponentially faster convergence rates during gradient descent. Basri *et al.* rigorously quantified this phenomenon by analyzing the convergence rate of neural networks for learned functions of different frequencies and proved that the time to reach a given error tolerance scales inversely with the neural tangent kernel eigenvalue associated with each frequency band [6].

Mathematically, consider a target function defined on a compact domain and expanded in an orthonormal Fourier basis with increasing frequencies. The neural tangent kernel admits an eigen-decomposition where the eigenvalues satisfy a rapid decay for higher frequencies for typical activation functions and architectures. During training, the residual error on each frequency component decays exponentially with a rate given by the corresponding kernel eigenvalue. Consequently, low-frequency components converge exponentially faster than high-frequency components. Xu *et al.* later extended this analysis to show that the bias persists even in very deep networks and is reinforced by the frequency principle across a wide range of architectures [9, 10].

Cao *et al.* further demonstrated that the spectral bias is not merely an initialization artifact but an intrinsic property of the optimization landscape under gradient flow [15]. In the infinite-width limit, the neural tangent kernel remains constant, and training reduces to kernel regression whose solution is provably biased toward the span of the leading eigenvectors.

Rahaman *et al.* corroborated these findings through spectral analysis on image data, showing that networks learn coarse global structure before fine local details [16].

A conceptual diagram of this phenomenon would display frequency on the horizontal axis (log scale from low to high) and learning rate on the vertical axis (log scale). The curve begins at a high plateau for low frequencies, then decays sharply after a cutoff frequency determined by network depth and width, approaching zero asymptotically for high frequencies. Superimposed would be training-time snapshots: after few epochs the low-frequency band is already near zero error, while high-frequency bands remain flat; only after orders-of-magnitude more training do the tails begin to descend. This visualization makes explicit why high-frequency targets require disproportionately long optimization and why early stopping or limited capacity exacerbates the bias.

When the target function is a force field rather than a scalar potential, the same bias applies component-wise. Because force is the negative gradient of the potential energy, high-frequency force components correspond to high-wavenumber derivatives, which inherit and amplify the spectral decay already present in the neural tangent kernel. This theoretical insight, grounded in the neural tangent kernel frequency response [6, 7], sets the stage for understanding why the bias becomes particularly severe once the input domain is replaced by the non-Euclidean geometry of a periodic crystal graph.

Graph Neural Networks on Periodic Crystals

Crystal Graph — A graph where nodes represent atoms, edges represent interatomic bonds (typically within a cutoff radius), with periodic boundary conditions.

Graph Laplacian — the degree matrix minus the adjacency matrix, whose eigenvalues characterize the graph's frequency components.

Periodic crystals impose a discrete translational symmetry that is faithfully captured by constructing a graph with periodic boundary conditions: each atom within the unit cell is a node, and edges connect atoms whose distance is less than a cutoff radius, with images replicated across cell boundaries. The resulting graph Laplacian is circulant in the periodic directions and therefore diagonalized by discrete Fourier modes whose wavevectors lie on the reciprocal lattice. Li *et al.* showed that the eigenvalues of the graph Laplacian directly control the frequency content of signals propagated by graph convolutions [11]. Low eigenvalues correspond to smooth, long-wavelength modes (acoustic phonons), while high eigenvalues encode rapid oscillations (optical branches near zone boundaries).

Message-passing neural networks, the dominant paradigm for crystal force prediction, implement a local aggregation rule that combines a node's own features with a weighted sum of its neighbors' features through learned functions. In the linear case, this reduces to multiplication by a normalized adjacency matrix whose eigenvalues lie in a fixed range and are strictly smaller than one in magnitude for non-trivial graphs. Repeated application therefore damps components associated with large Laplacian eigenvalues, functioning as a low-pass filter. Schütt *et al.* employed precisely this message-passing construction in SchNet to achieve state-of-the-art force prediction on molecular and crystalline systems [2], while Xie *et al.* extended it to crystal graphs with explicit periodic encoding in crystal graph convolutional neural networks [3]. Batzner *et al.* later introduced E(3)-equivariant layers that preserve rotational symmetry but retain the same underlying low-pass message-passing structure [4].

The relationship between graph frequency and physical vibrational frequency is direct: the dynamical matrix for phonons is closely related to the Hessian of the potential, which itself is the second derivative of the learned energy; hence the eigenvectors of the graph Laplacian approximate the normal modes when the cutoff graph captures the relevant interaction range. High-frequency optical phonons therefore project onto high-Laplacian-eigenvalue subspaces. Because standard graph neural network layers attenuate these subspaces, the predicted force field necessarily under-represents short-wavelength vibrations.

Gilmer *et al.* established the message-passing framework as a general template for quantum chemistry [17], and subsequent works such as PhysNet [12] and atomistic line graph neural networks [13] confirmed that increasing depth or introducing hierarchical pooling further reinforces the low-pass character. Chen *et al.* analyzed simple deep graph convolutional networks and demonstrated that their effective frequency response is governed by repeated powers of the normalized Laplacian [18]. When these networks are applied to periodic crystals, the discrete translational symmetry guarantees that the Laplacian spectrum contains gaps whose size scales with the inverse square of the unit-cell size; larger cells produce denser but still gapped spectra, while small primitive cells produce large spectral gaps that separate acoustic from optical branches even more sharply.

Thus, the combination of periodic boundary conditions and message-passing induces a structural bias against high-frequency force components that is absent in non-periodic molecular graphs and molecular machine-learning benchmarks [19] or in grid-based image data where convolutional filters can be designed with arbitrary frequency response. This structural amplification of spectral bias, absent from the original neural tangent kernel analyses [6, 16], is the central reason why force prediction on crystals exhibits the observed frequency-dependent error pattern.

Why High-Frequency Forces are Theoretically Hard

For a periodic crystal graph, the high-frequency components of the force field lie in the eigenspaces of the graph Laplacian with large eigenvalues.

To see this, expand the target force vector field in the orthonormal eigenbasis of the crystal graph Laplacian. Physical intuition and the dynamical matrix confirm that short-wavelength phonon modes—precisely those carrying high-frequency force information—correspond to eigenvectors with large eigenvalues near the upper end of the spectrum. Thus high-frequency forces are synonymous with projections onto the high-eigenvalue tail.

Standard message-passing graph neural networks apply a low-pass filter to node features, attenuating components corresponding to large Laplacian eigenvalues.

Consider a single linear message-passing step with the normalized adjacency operator. In the Laplacian eigenbasis, this multiplies each coefficient by a factor strictly less than one for high-frequency modes. After many layers the attenuation factor becomes very small for those modes, which for large eigenvalues decays exponentially in depth. Oono and Suzuki proved that graph neural networks exponentially lose expressive power precisely because of this eigenvalue damping [20], while Chen *et al.* quantified the over-smoothing effect in deep graph convolutional networks [21]. Applied to force prediction, the predicted force therefore systematically underestimates high-frequency contributions.

The neural tangent kernel of a graph neural network on a periodic crystal graph has eigenvalues that decay with graph Laplacian frequency, causing slower learning for high-frequency force components.

In the infinite-width limit, the graph neural network neural tangent kernel is a kernel on the product space of node features and graph structure. Because the message-passing operator commutes with the Laplacian eigenbasis, the neural tangent kernel inherits the same eigenbasis, and its eigenvalues decrease as the Laplacian eigenvalue increases. Xu *et al.* showed that the neural tangent kernel eigenvalues of graph networks decay with the graph frequency index [22], and the periodic crystal case inherits an additional discrete gap structure. Consequently, gradient descent on the high-frequency force residuals converges at a much slower rate, which is orders of magnitude slower than for low-frequency residuals.

Combining Propositions yields the central theoretical claim: high-frequency forces are theoretically hard because they live in the high-eigenvalue subspace of the crystal graph, message passing damps that subspace exponentially, and the neural tangent kernel itself assigns vanishingly small learning rates to those directions. The remainder of the paper formalizes these intuitions into a complete framework and derives quantitative bounds.

A Theoretical Framework for Spectral Failure

Frequency decomposition

Decompose the target force field into low-frequency and high-frequency components using the graph Laplacian eigenbasis. Choose a spectral cutoff separating acoustic from optical branches; the low-frequency part collects all components below the cutoff and the high-frequency part collects all components above it.

Learning dynamics

Model the training dynamics of a graph neural network for force prediction using neural tangent kernel theory. In the infinite-width regime the parameter update follows gradient flow under the neural tangent kernel. Projecting onto the Laplacian eigenbasis diagonalizes both the loss and the kernel, decoupling the dynamics for each frequency mode.

Error bound

The high-frequency error after a given number of training steps remains bounded below by an exponentially decaying term whose rate is controlled by the smallest high-frequency neural tangent kernel eigenvalue, which itself is limited by the graph Laplacian spectral gap.

Periodicity amplification

Periodic boundary conditions create a discrete spectrum with gaps, making high-frequency components harder to learn than in non-periodic graphs. In a finite molecular graph the Laplacian eigenvalues are dense near zero and the spectral gap is small; in a crystal the Brillouin-zone sampling enforces larger gaps, shifting high-frequency modes farther into the damped tail of the neural tangent kernel.

Taken together, these four components constitute a closed theoretical account: the force field is spectrally decomposed, the graph neural network learns each component at a rate dictated by the neural tangent kernel eigenvalue inherited from the Laplacian, high-frequency modes receive exponentially small rates and therefore retain large residuals, and periodicity enlarges the effective spectral gap, amplifying the entire mechanism. The framework explains not only the existence of the bias but its quantitative severity in crystalline versus molecular settings.

Mathematical Mechanisms

Repeated message passing averages node features, attenuating high-frequency components exponentially with depth. In the eigenbasis, each layer multiplies coefficients by a damping factor less than one for high-frequency modes; after many layers the high-frequency amplitude decays rapidly toward zero. When predicting forces, this directly suppresses short-wavelength variations, producing overly smooth force fields. Li *et al.* first identified this smoothing in graph convolutional networks [11], and later analyses quantified the exponential loss of information [21, 23]. The gap between low and high eigenvalues of the graph Laplacian creates a separation in neural tangent kernel eigenvalue magnitudes. The gap induces an exponential suppression of learning rates for high-frequency modes, absent in Euclidean convolutional networks. This explains why crystal graph neural networks require far more epochs to fit optical phonons than acoustic modes.

Global or hierarchical pooling operations act as low-pass filters, discarding high-frequency information. In crystal force predictors that employ readout or hierarchical coarsening, the pooling step projects onto the constant eigenvector subspace, annihilating all components orthogonal to it. The residual high-frequency content is therefore lost irreversibly, independent of depth. Each message-passing step only aggregates local neighbors, limiting the receptive field for short-wavelength high-frequency forces. To resolve a short wavelength, the receptive field must span a sufficient distance; yet after a fixed number of layers the effective receptive field grows only linearly, while high-frequency phonons require

exponentially larger fields in periodic cells. Xu *et al.* proved that graph neural network expressive power is bounded by the 1-WL test precisely because of this locality [24], which in the force-prediction setting translates into an inability to capture rapidly oscillating force patterns. **Table 2** differentiates the four suppression mechanisms analytically, showing that the manuscript’s explanation depends on structurally distinct but mutually reinforcing sources of high-frequency loss.

Table 2. Comparative Mechanism Matrix for Spectral Suppression of High-Frequency Forces in Periodic-Crystal GNNs

Mechanism	Formal driver	Spectral action on high-frequency modes	Distinctive signature in force prediction	Why periodic crystals intensify it	Primary architectural vulnerability
Over-smoothing	Repeated neighbor aggregation through stacked message-passing layers	Successive attenuation of large-eigenvalue components toward smooth node representations	Predicted force fields become overly smooth and under-responsive to short-wavelength variations	Crystal graphs place physically meaningful phonon information directly in the suppressed spectral tail	Increasing depth without protected high-frequency channels
Spectral-gap bottleneck	Separation between low- and high-eigenvalue graph subspaces	Learning-rate disparity widens between low- and high-frequency residuals	Optical or boundary-zone vibrational components remain poorly fitted even when low-frequency forces converge	Discrete periodic spectra sharpen acoustic–optical separation more than in non-periodic graphs	Architectures relying entirely on inherited graph filtering and unmodified NTK ordering
Pooling-induced low-pass filtering	Global readout or hierarchical coarsening	Projection toward coarse or smooth signal content with loss of oscillatory detail	Fine-scale force variations disappear at intermediate or output stages	Periodic crystal signals often require retention of structured local oscillation across repeated cells	Readout schemes that compress before preserving high-frequency information
Locality of message passing	Finite receptive field growth under local aggregation	Short-wavelength interactions are not fully resolved across the spatial range needed for rapid oscillation patterns	Local force directions may appear plausible while phase-sensitive high-frequency structure is missed	Repeated unit-cell structure can require coordinated representation beyond the effective local neighborhood	Shallow local architectures without multiscale or nonlocal correction pathways
Mechanism interaction effect	Joint action of damping, rate suppression, compression, and locality	High-frequency content is not merely learned slowly but progressively weakened in both representation and optimization	Error concentrates in the physically consequential tail rather than being uniformly distributed	Periodicity aligns physical importance with spectral disadvantage	Conventional sequential message-passing pipelines

These mechanisms operate synergistically: over-smoothing and locality damp the signal at each layer, the spectral gap reduces the learning rate, and pooling removes whatever remains. The net effect is a systematic, theoretically

unavoidable bias against high-frequency forces in any standard message-passing graph neural network applied to periodic crystals.

Bounds and Limitations

For a graph neural network with L layers and a message-passing operator whose eigenvalues are denoted μ_k , the convergence rate for any frequency component k is proportional to μ_k^L raised to the power L . This means that high-frequency components, which carry eigenvalues close to the upper end of the spectrum, experience an exponentially decaying learning speed as the network depth increases. In practice, this bound shows that even moderate depth quickly renders high-frequency force components nearly unlearnable within realistic training budgets.

Under standard assumptions of infinite network width and gradient-flow training, the high-frequency force prediction error cannot fall below a positive threshold determined by the spectral gap of the crystal graph Laplacian. Specifically, once the low-frequency components have been fitted, the residual error on short-wavelength forces remains bounded away from zero by a term that scales inversely with the smallest high-frequency neural tangent kernel eigenvalue. This lower bound arises directly from the ordering of learning rates established in the frequency principle and is amplified by the discrete gaps inherent to periodic structures.

Increasing network depth improves accuracy on low-frequency forces but simultaneously degrades accuracy on high-frequency forces because of progressive over-smoothing. The trade-off is quantitative: each additional layer multiplies the high-frequency attenuation factor, producing a monotonic rise in high-frequency error even as overall training loss continues to decrease.

These bounds rest on three key assumptions. First, the network is assumed to operate in the infinite-width regime where the neural tangent kernel remains constant throughout training. Second, the crystal graph is treated as fixed with periodic boundary conditions, so the Laplacian spectrum is static and gapped. Third, the force field is expanded exactly in the Laplacian eigenbasis, an idealization that holds when the cutoff radius captures the dominant interactions. Relaxing any of these assumptions softens the bounds but does not eliminate the qualitative spectral bias. For finite-width networks, additional variance terms appear in the error, yet the leading exponential decay with frequency persists. When the cutoff radius is chosen too small, the graph becomes disconnected and the spectral gap widens further, strengthening the bounds against high-frequency learning.

The limitations of these bounds are equally important, particularly when force prediction is coupled to larger multiscale simulation settings [25]. They apply most strictly to linear or mildly nonlinear message-passing layers; highly nonlinear activations can introduce weak high-frequency pathways, but these pathways remain sub-dominant and do not overturn the overall bias. The bounds also assume mean-squared-error loss on forces; other loss functionals that explicitly weight high-frequency components could shift the effective neural tangent kernel eigenvalues but cannot remove the underlying spectral ordering. Finally, the analysis is purely theoretical and does not quantify the exact numerical value of the spectral gap for any particular material; that value must be computed from the specific crystal graph. Nevertheless, the bounds provide a universal explanation for the observed frequency-dependent errors across the entire class of periodic-crystal force predictors and establish a clear ceiling on what conventional architectures can achieve.

Relation to Other Theoretical Results

The spectral bias identified here connects directly to the expressive power of graph neural networks as captured by the Weisfeiler-Lehman hierarchy. Xu *et al.* demonstrated that standard message-passing networks are at most as powerful as the 1-WL test, which relies on iterative local aggregation and therefore cannot distinguish graphs beyond local neighborhoods. In the context of periodic crystals, this limited expressivity translates into an inability to capture short-

wavelength force patterns that require global coordination across multiple unit cells. The same locality that limits graph isomorphism testing also limits high-frequency force reconstruction.

Our framework also deepens the understanding of over-smoothing and over-squashing. Li *et al.* first highlighted how repeated graph convolutions drive node features toward a constant vector, erasing high-frequency information. Chen *et al.* later quantified this smoothing through the powers of the normalized Laplacian, showing exponential decay of non-constant eigenvectors. The present analysis shows that over-smoothing is not merely a numerical artifact but the direct consequence of the neural tangent kernel inheriting the Laplacian spectrum. Over-squashing, in which distant node information is compressed into fixed-size vectors, further exacerbates the problem for high-frequency forces that oscillate rapidly across the crystal lattice.

The spectral failure also relates to graph isomorphism and symmetry considerations. Periodic boundary conditions impose additional translational symmetries that discretize the Laplacian spectrum more sharply than in finite molecular graphs. This discretization creates larger spectral gaps, which in turn widen the separation of neural tangent kernel eigenvalues. Consequently, the same symmetry that makes crystal graphs computationally convenient also makes high-frequency learning theoretically harder. Approximation theory for graph neural networks provides another link: previous results on universal approximation hold only for continuous functions on compact domains, yet they do not guarantee uniform approximation rates across frequency bands. The frequency-dependent bounds derived here supply the missing rate information and explain why approximation guarantees appear strong for energies but weaken dramatically for forces.

Taken together, these connections demonstrate that spectral bias is not an isolated phenomenon but the intersection point of several well-established theoretical limitations in graph learning. It unifies the frequency principle from classical neural networks with the spectral and topological constraints unique to graph-structured data on periodic domains.

Implications for Model Design

Deeper graph neural networks worsen spectral bias. Therefore, architectures should favor shallower networks or incorporate residual connections that explicitly preserve high-frequency pathways. By adding skip connections around message-passing layers, designers can prevent the cumulative low-pass filtering that would otherwise suppress short-wavelength forces. Equivariant networks and equivariant message-passing approaches for tensorial properties [26] may partially mitigate spectral bias through better inductive biases. The E(3)-equivariant layers introduced by Batzner *et al.* already encode rotational symmetries that align more closely with physical force fields. Extending these layers with explicit high-frequency basis functions could further reduce the reliance on learned message passing and thereby lessen the inherited Laplacian damping. Multi-scale architectures such as graph U-nets can preserve high-frequency information through skip connections. By routing high-eigenvalue features directly from early to late layers, these models bypass the repeated attenuation that occurs in purely sequential message passing. The skip pathways act as frequency-preserving highways, allowing high-frequency force components to reach the output without being filtered out.

Fourier feature embeddings and adaptive frequency-response filtering [27] can pre-encode high-frequency components before message passing begins. By lifting node and edge features into a higher-dimensional Fourier basis that explicitly includes short-wavelength modes, the network receives a richer initial representation that counteracts the low-pass nature of subsequent layers. This preprocessing step shifts part of the spectral burden from learning to architecture design. Spectral methods that directly incorporate the graph Fourier transform may become necessary for accurate high-frequency force prediction. Instead of relying solely on local message passing, future models could perform explicit filtering in the Laplacian eigenbasis, applying learned corrections only to high-frequency modes while keeping low-frequency modes stable. Such hybrid spectral-graph networks would combine the strengths of convolutional filtering with the frequency control offered by Fourier analysis.

Table 3 converts the manuscript's theoretical claims into an explicit design framework, showing how each identified source of spectral failure implies a distinct architectural intervention.

Table 3. Theory-to-Design Translation Framework for Mitigating Spectral Bias in Periodic-Crystal Force Predictors

Theoretical problem identified in the manuscript	Why standard GNNs fail	Design principle implied by the theory	Candidate architectural response	Expected benefit	Trade-off or caution
High-frequency modes receive weak learning rates under NTK dynamics	Optimization favors low-frequency residuals first and disproportionately	Increase direct accessibility of high-frequency information	Residual high-frequency pathways or skip routes that bypass repeated damping	Preserves short-wavelength signal during training and inference	Can reduce simplicity and may require frequency-aware regularization
Message passing acts as a repeated low-pass filter	Layer stacking progressively smooths large-eigenvalue components	Protect high-frequency content from cumulative attenuation	Shallow message-passing backbones with structured residual connections	Reduces over-smoothing while retaining local physical inductive bias	Too little depth may weaken low-frequency relational modeling
Periodic crystal graphs sharpen spectral gaps	High-frequency subspaces become more isolated and harder to learn	Counteract gap-induced suppression before or during propagation	Spectral preconditioning or graph-Fourier-informed correction modules	Improves access to modes disadvantaged by graph structure	Requires additional spectral computation and graph-specific design choices
Local aggregation cannot fully resolve rapidly oscillating patterns	Effective receptive field grows too slowly relative to the structure of short-wavelength modes	Introduce explicit multiscale or nonlocal pathways	Graph U-net variants, multiresolution pathways, or hybrid local-global architectures	Better reconstruction of oscillatory force structure across spatial scales	Added complexity may complicate training stability and interpretation
Force targets contain physically meaningful short-wavelength content from the outset	Raw feature spaces may not expose these modes sufficiently	Encode frequency diversity before message passing begins	Fourier feature embeddings for node or edge descriptors	Makes high-frequency structure more visible to downstream layers	Poorly chosen encodings may inject irrelevant oscillatory content
Standard architectures lack controllable frequency selectivity	The model inherits frequency response indirectly rather than designing it explicitly	Move from implicit to explicit spectral control	Hybrid spectral-graph architectures with learned correction in Laplacian space	Enables targeted improvement of the high-frequency tail without sacrificing low-frequency stability	Greater mathematical and computational burden

These implications are not speculative; they follow rigorously from the derived bounds and mechanisms. Any architecture that fails to address the spectral gap and the low-pass character of message passing will inherit the same high-frequency limitations identified throughout this analysis [28]. The theoretical framework therefore supplies a clear roadmap for next-generation force-field models that maintain accuracy across the entire phonon spectrum.

Conclusion

This theoretical analysis has established that graph neural networks for periodic crystals exhibit a pronounced spectral bias: they preferentially learn low-frequency components of the target force field while systematically missing high-frequency components. The central problem arises because high-frequency forces live in the high-eigenvalue subspace of the crystal graph Laplacian, standard message-passing layers act as repeated low-pass filters, and the neural tangent kernel assigns exponentially smaller learning rates to those high-frequency directions. Periodic boundary conditions further amplify the difficulty by introducing discrete spectral gaps that separate acoustic and optical branches more sharply than in molecular graphs.

We formalized spectral bias in the graph setting, defined the crystal graph and its Laplacian spectrum, articulated three informal propositions that explain the theoretical hardness of high-frequency forces, and assembled a four-component framework that decomposes the force field, models training dynamics via the neural tangent kernel, derives error bounds, and accounts for periodicity amplification. Four mathematical mechanisms—over-smoothing, spectral-gap bottlenecks, pooling-induced filtering, and locality of message passing—were shown to operate synergistically, each contributing to the suppression of short-wavelength information. Explicit convergence-rate bounds and high-frequency error lower bounds were presented, together with a depth-frequency trade-off that quantifies the fundamental limitations of conventional architectures. These results were related to broader theoretical literature on expressive power, over-smoothing, and approximation theory, confirming that the observed bias is the natural intersection of known graph-learning constraints. Finally, five concrete implications for model design were offered, pointing toward residual connections, equivariant enhancements, multi-scale pathways, Fourier embeddings, and hybrid spectral methods as promising routes forward.

The framework and bounds supplied here close a critical gap in the theoretical understanding of machine-learned force fields. They explain, without recourse to any empirical benchmark, why high-frequency forces corresponding to short-wavelength phonons are systematically missed and why this omission matters for thermal transport, phonon dispersion, and mechanical stability. Future architecture development can now be guided by spectral considerations rather than trial-and-error alone. By incorporating explicit high-frequency preservation mechanisms, the next generation of graph neural networks for materials science can achieve uniformly accurate force predictions across all relevant frequency bands, unlocking more reliable simulations of dynamic and anharmonic phenomena in crystalline materials.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 28 Apr 2021

Revised: 28 Aug 2021

Accepted: 15 Nov 2021

Published online: 18 January 2022

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

Gurney K. An introduction to neural networks. Boca Raton: CRC Press; 2018.

Schütt KT, Sauceda HE, Kindermans PJ, Tkatchenko A, Müller KR. SchNet: A deep learning architecture for molecules and materials. *J Chem Phys.* 2018;148(24).

Xie T, Grossman JC. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys Rev Lett.* 2018;120(14):145301.

Wu Z, Pan S, Chen F, Long G, Zhang C, Yu PS. A comprehensive survey on graph neural networks. *IEEE Trans Neural Netw Learn Syst.* 2021;32(1):4-24.

Ingman VM, Schaefer AJ, Andreola LR, Wheeler SE. QChASM: Quantum chemistry automation and structure manipulation. *Wiley Interdiscip Rev Comput Mol Sci.* 2021;11(4):e1510.

Ronen B, Jacobs D, Kasten Y, Kritchman S. The convergence rate of neural networks for learned functions of different frequencies. *Adv Neural Inf Process Syst.* 2019;32.

Xu ZQ, Zhang Y, Luo T, Xiao Y, Ma Z. Frequency principle: Fourier analysis sheds light on deep neural networks. *arXiv preprint arXiv:1901.06523.* 2019 Jan 19.

Suzuki T, Tamura R, Miyazaki T. Machine learning for atomic forces in a crystalline solid: Transferability to various temperatures. *Int J Quantum Chem.* 2017;117(1):33-9.

Xu ZJ, Zhou H. Deep frequency principle towards understanding why deeper learning is faster. In: *Proc AAAI Conf Artif Intell.* 2021;35(12):10541-50.

Luo T, Ma Z, Xu ZQ, Zhang Y. Theory of the frequency principle for general deep neural networks. *arXiv preprint arXiv:1906.09235.* 2019 Jun 21.

Li Q, Han Z, Wu XM. Deeper insights into graph convolutional networks for semi-supervised learning. In: *Proc AAAI Conf Artif Intell.* 2018;32(1).

Unke OT, Muwly M. PhysNet: A neural network for predicting energies, forces, dipole moments, and partial charges. *J Chem Theory Comput.* 2019;15(6):3678-93.

Choudhary K, DeCost B. Atomistic line graph neural network for improved materials property predictions. *npj Comput Mater.* 2021;7(1):185.

Babaei H, Guo R, Hashemi A, Lee S. Machine-learning-based interatomic potential for phonon transport in perfect crystalline Si and crystalline Si with vacancies. *Phys Rev Mater.* 2019;3(7):074603.

Cao Y, Fang Z, Wu Y, Zhou DX, Gu Q. Towards understanding the spectral bias of deep learning. *arXiv preprint arXiv:1912.01198.* 2019 Dec 3.

Rahaman N, Baratin A, Arpit D, Draxler F, Lin M, Hamprecht F, et al. On the spectral bias of neural networks. In: *Int Conf Mach Learn.* 2019. p. 5301-10.

Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE. Neural message passing for quantum chemistry. In: *Int Conf Mach Learn;* 2017. p. 1263-72.

Chen M, Wei Z, Huang Z, Ding B, Li Y. Simple and deep graph convolutional networks. In: *Int Conf Mach Learn;* 2020. p. 1725-35.

Wu Z, Ramsundar B, Feinberg EN, Gomes J, Geniesse C, Pappu AS, et al. MoleculeNet: A benchmark for molecular machine learning. *Chem Sci.* 2018;9(2):513-30.

Oono K, Suzuki T. Graph neural networks exponentially lose expressive power for node classification. *arXiv preprint arXiv:1905.10947.* 2019 May 27.

Chen D, Lin Y, Li W, Li P, Zhou J, Sun X. Measuring and relieving the over-smoothing problem for graph neural networks from the topological view. In: *Proc AAAI Conf Artif Intell.* 2020;34(04):3438-45.

Xu K, Li J, Zhang M, Du SS, Kawarabayashi KI, Jegelka S. What can neural networks reason about? *arXiv preprint arXiv:1905.13211.* 2019 May 30.

Cai C, Wang Y. A note on over-smoothing for graph neural networks. *arXiv preprint arXiv:2006.13318.* 2020 Jun 23.

Xu K, Hu W, Leskovec J, Jegelka S. How powerful are graph neural networks? *arXiv preprint arXiv:1810.00826.* 2018 Oct 1.

Veske M, Kyritsakis A, Eimre K, Zadin V, Aabloo A, Djurabekova F. Dynamic coupling of a finite element solver to large-scale atomistic simulations. *J Comput Phys.* 2018;367:279-94.

Schütt K, Unke O, Gastegger M. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In: *Int Conf Mach Learn;* 2021. p. 9377-88.

Dong Y, Ding K, Jalaian B, Ji S, Li J. AdaGNN: Graph neural networks with adaptive frequency response filter. In: *Proc 30th ACM Int Conf Inf Knowl Manag;* 2021. p. 392-401.

Hu W, Fey M, Zitnik M, Dong Y, Ren H, Liu B, et al. Open graph benchmark: Datasets for machine learning on graphs. *Adv Neural Inf Process Syst.* 2020;33:22118-33.